Numerical model estimation of Biomethane production using an anaerobic CSTR: model formulation, parameter estimation and uncertainty/sensitivity analysis

Yazidi, Hatem; Thanikal, Joseph V. ; Bevan, Geraint; Pahl, Ole; Hunter, Colin

Publication date:
2018

Document Version
Author accepted manuscript

Numerical model estimation of Biomethane production using an anaerobic CSTR: model formulation, parameter estimation and uncertainty/sensitivity analysis

Hatem Yazidi,
Joseph V. Thanikal,
Dept. of Built and Natural Environment
Caledonian College of Engineering
P.O.Box : 2322 CPO Seeb 111
Muscat, Sultanate of Oman
E-mail: hatem@caledonian.edu.om,
and jthanikal@gmail.com

Geraint Bevan, Ole Pahl,
and Colin Hunter
School of Engineering and Built Environment
Glasgow Caledonian University
Cowcaddens Rd, Glasgow G4 0BA, UK
E-mail: Geraint.Bevan@gcu.ac.uk, O.Pahl@gcu.ac.uk and
Colin.Hunter@gcu.ac.uk

Abstract— **There is an essential need to find an effective and reliable technique to transform organic waste into energy. The increasing knowledge of anaerobic processes as a reliable solution to transform organic waste into energy has led to the development and implementation of sophisticated and complex mathematical models. Quantitatively speaking, on 2017, the estimated generation of solid waste in the Sultanate of Oman amounts to about 1.7 million tonnes per annum (1.2 kg/capita/day). Those quantities are mostly directly dumped with no or little treatment. Despite the continuous efforts conducted by the local government, currently, there is still a significant volume of untreated solid waste material being sent to landfills, thus adding to the atmosphere, emissions of thousands of metric tonnes per annum of methane and carbon dioxide, the most adverse greenhouse gases. Therefore, deploying mechanisms to control organic waste dumping process is counted as one of the country priorities. Consequently, estimating in advance the biogas quantity of a given organic waste raw material become primordial for municipal solid waste managers. The latter can be done by using numerical models that have the ability to describe the anaerobic demanding processes. However, there is little consensus about the model's structure and parametric identifiability questions. Those questions are not yet sufficiently elucidated in the reported anaerobic digestion modelling studies. In this paper, a complex numerical model is proposed for simulating the anaerobic biogas production from the co-digestion of organic waste material. An innovative complex dynamic model structure is proposed to support full-scale anaerobic plant design and operation decisions and to assist laboratory scale and pilot co-digestion research. The model facilitates the understanding of the co-digestion effects and therefore discards any potential negative impacts from mixing based on random or heuristic decisions. This paper introduces an innovative modelling procedure, including the application of uncertainty and global sensitivity analysis (LHS/PRCC/eFAST), which allows the study of a multi-dimensional parameter space globally so all uncertainties can be identified among the parameters; a multi-steps approach that gives a clear picture of the main sensitive model parameter. Among them, special concerns will be given to those identified as sensitive conducting to the digester failure. Sophisticated and stable algorithms are designed for the model cost function minimization criteria and result in an increase of the model accuracy. The model parameter uncertainty and sensitivity analysis revealed that the hydrolysis and acidogenesis phases are the most affecting steps of the methane production. A parameter such the polymer hydrolysis rate, the specific acidogens maximum growth rate, the saturation constant for acidogens, the specific acetoclastic methanogens maximum growth rate, the saturation constant for acetoclastic methanogens, and the gas-liquid mass transfer coefficient for CH4, contribute the most to the variance of the complex model estimate of methane.**

Keywords—Latin hypercube sampling (LHS), Partial rank correlation (PRCC), Extended Fourier amplitude sensitivity test (eFAST), Monte-Carlo methods (MC), Global Sensitivity Analysis

I. INTRODUCTION

The main Anaerobic co-Digestion (AcoD) process outputs is the production of biogas. The latter mechanism can be seen as a clean and sustainable process for producing electricity through biogas combustion. The AcoD process is orchestrated by a consortium of microorganisms that degrade organic substrates present in the biological wastes. AcoD can, also, have a key role in mitigating the adverse effect of uncontrolled dumping all types of solid waste by transforming the organic fraction of it into fertilizer [1], [2]. Nevertheless, from the practical point of view, still, some problems arising from the industrialisation of AcoD process are not resolved yet. They generally are the same as those encountered in any processing industry, mainly controlling the biotechnological mechanisms of an AcoD plant to avoid any product failure that evolves delays in the final delivery. In Generally, AcoD plant

designers and operators actually seek for optimisation of the degree of the initial substrates with inoculum mixing to lower the cost and, also, the environmental impact without compromising the biogas output [3]. The solution to the previous advanced situation depend closely on and can be influenced by, the substrate type as well as by the process flow dynamics in the AcoD plant digesters. The second is in turn determined by the physical parameters of the digestion vessels, inflow mode, sludge rheology and, crucially mixing systems. Limiting steps to AcoD can be considered fatal to the total system and includes the acetogenesis mechanism [4]-[6], methanogenesis mechanism [7], hydrolysis mechanism [8] and disintegration mechanism [9], [10].

This paper introduces a useful numerical tool that reduces considerably the uncertainty about AcoD mechanisms: a complex deterministic dynamic mathematical model, adapting the operational conditions to simulate the co-digestion processes that can indicate the best settings to maximise the performance of the AcoD process.

To deal with model inputs uncertainty we propose in this manuscript an uncertainty analysis (UA) method to investigate the model output uncertainty that is generated from uncertainty in parameter inputs. Sensitivity analysis (SA) simply constitute a complementary analysis to UA as it assesses how variations in model outputs can be apportioned, qualitatively or quantitatively, to different input sources.

II. MATERIAL AND METHODS

A. Substrates

The fruit and vegetable waste were used in this study and were collected from Al Mawalah Central Market in Muscat. The cocked oil was collected from nearby restaurants. All solid substrates were shredded into small pieces and stored at 4 degrees Celsius and characterised for total solids (TS), Suspended solids (SS) and Volatile suspended solids (VS) as detailed in Table 1. The substrates were characterised as per APHA (2012) methods [11]. The cooked oil contains rich amounts of lipids. At higher concentration lipids are considered to be problematic components for better performance of an AcoD process [12]. Cooked oil is often co-digested with other substrate types to reduce the lipid concentration in the digester [13]. Consequently, problems such clogging, adsorption to biomass (affecting the mass transfer process), microbial inhibition due to the degradation could be a trigger for enhancing the long-chain fatty acids (LCFA). An imminent byproduct of the lipids transformation which increase the digetsers acidification process and, therefore, increase the chances to bioreactor collapsing.

B. Reactor design and operation

A multi-series of experiments were performed in an identical two double-walled bio-reactors of 6 litres effective volume (BR) maintained at 38°C by a regulated water bath (Fig. 1). Mixing in the BRs was done by a system of magnetic stirring. The pH inside the reactor was continuously monitored online using Metler Toledo pH probe Inpro 4260i and maintained at 7.5±0.5.

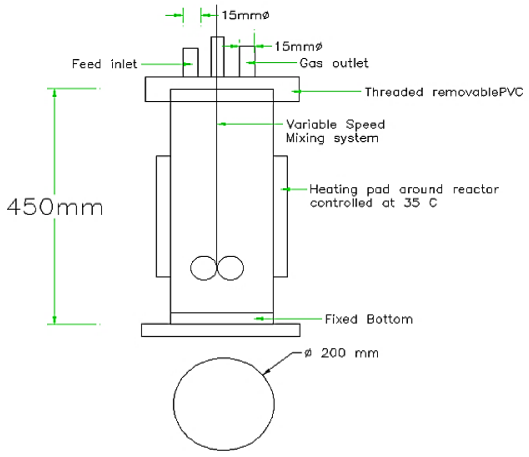


Fig 1. Schematic diagram of the laboratory-scale digester

Parameters	Potato	Carrot	Spinach	Onion	Tomato	Lettuce	Apple	Orange	Grapes	POM*	WAT**
PH	4.5	4.8	5.6	7.82	7.78	7.70	5.74	3.89	3.42	5.31	6.79
COD(g/Kg)	220	180	90	52	310	145	120	60	80	210	150
MC†(%)	99.36	89.69	99.7	85.56	99.5	88.54	90.47	95.92	98.8	72.63	95.35
TS††(g/g)	0.34	0.17	0.13	0.25	0.09	0.11	0.21	0.25	0.36	0.38	0.14
VS‡(g/g)	0.16	0.07	0.05	0.12	0.04	0.05	0.12	0.12	0.16	0.20	0.08
SS‡‡(g/g)	0.18	0.09	0.08	0.13	0.05	0.06	0.13	0.14	0.20	0.25	0.09

TABLE 1. SUBSTRATES CHARACTERISTICS OF AND INOCULUM USED IN THE EXPERIMENTS
*POM: Pomegranate, **WAT: Watermelon, † MC:Moisture Content, † † TS: Total Solids, ‡ VS: Volatile Solids and ‡ ‡ SS: Suspended Solids.

This project was funded by the Sultanate of Oman research council (TRC – Grant 63) .

The BR was operated in batch mode without withdrawal (semi-continuous). The flow rate was determined by each batch assay. The batch end was considered once the flow rate reached a threshold value of 1 ml/h. The reactor was fed with vegetable substrates at an OLR varying from 1.0 to 6.0 g [VS]/l, respectively.

C. *Inoculum*

A quantity of 700 g of Granular sludge obtained from an Up-Flow Anaerobic Sludge Blanket Up-flow Anaerobic Sludge fixed-Bed (UASB) reactor treating sugar factory effluent was used to inoculate our 6l volume bioreactor. In addition, we fed our BR with ethanol during the startup phase of our experiments to observe the biological activity of the inoculum.

D. *The experimental data*

The total of 120 days of cumulative biogas production temporal series plot is shown in Fig. 2, where the biogas production is observed to be exponential to the increase in Organic Loading Rate (OLR). Nevertheless, the reactor experienced the problem of mixing when higher quantities of solids are added to the reactor. It is worth noting that some failures such leaks, tube clogging and so forth have slightly perturbed the initial protocol. Fig. 3 displays some of the batches conducted for fruits co-digestion.

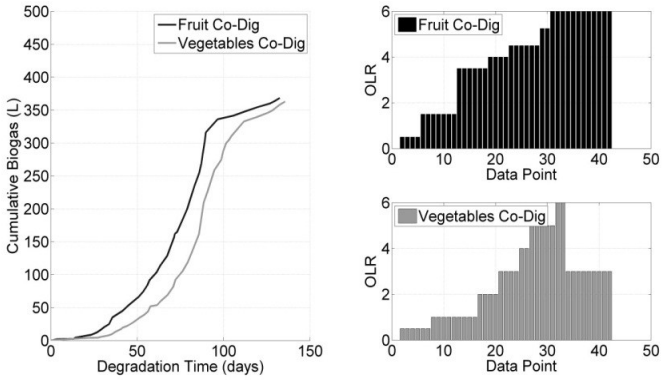


Fig 2. Cumulative biogas production (left plot) Vs. organic loading rate for fruit, vegetables and cooked oil waste (right plots).

III. MODEL ASSUMPTION AND DESCRIPTION

A. *Model assumptions*

The selection process of any dynamic model shall be based on the trade-off between the model complexity, flexibility and avidity (determined by the number of state variables and parameters included). In this paper, we present a model that is data-driven and mechanistically describing digestion dynamic processes. The selected model is partially driven by the amount of a priori knowledge available on the system: a combination of multiple parameters such as measured data, bioreactor design, yields coefficients, bacterial growth rates, substrates initial concentrations and error estimation.

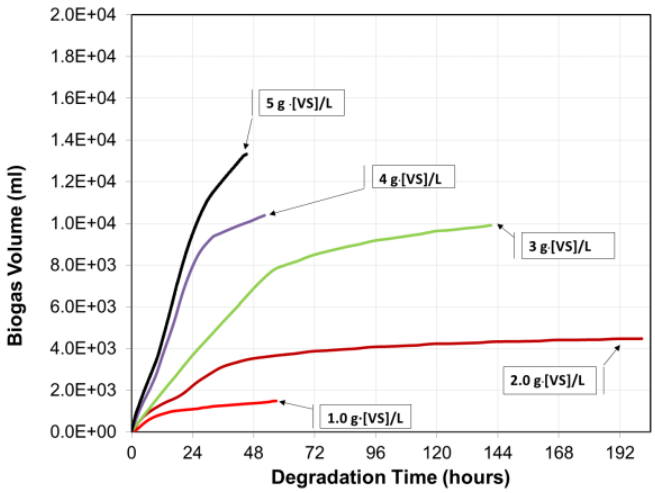


Fig 3. Fruits co-digestion cumulative biogas production tests

Furthermore, the proposed mechanistically model parameters have a physical interpretation, but are adjustable, for instance, by a parameter estimation procedure. Although mathematical models are efficient tools used to optimise an AcoD process, it is worth considering that obtaining reliable parameters of an anaerobic digestion is very challenging. In fact, the AcoD processes compass a very wide range of microorganisms and compounds. Nevertheless, in this work we assume that the bacterial populations can be divided into five main groups of homogeneous characteristics, and that the proposed model can be described by the seven-stage process as described in the next Section here below. In addition, we assume that the reactor behaves like a perfectly mixed tank and that the biomass is uniformly distributed inside the fermentor.

B. *Model description*

The reactions and processes occurring in the anaerobic digestion proposed system to model are simplified into the following general steps (Fig. 4). As shown in Fig. 4, the proposed model structure is typically composed of the combination of hydrodynamics terms, liquid−gas terms, and conversion (kinetic + yields) terms. The conversion and liquid−gas transfer terms contain all the parameters to be calibrated, while the terms related to the hydrodynamics are ideally characterised by the known values of the influent. at the first step, the polymeric substrate (S0) is hydrolysed using enzymes, producing fermentable monomers (S1). Then, at the second step, the fermentable monomers (S1) are transformed into Propionic acid (S2), soluble hydrogen (S3), soluble carbon dioxide (S4), Acetic acid (S5) and Butyric acid (S6) by acidogenic bacteria (X1). The third step is characterised by the Propionic acid (S2) transformation into H2 (S3), followed by its transformation into CO2 (S4) and Acetic acid (S5) by syntrophic bacteria A (X2). The fourth step is resumed by the transformation of the Acetic acid (S5) into methane (S7) and CO2 (S4) by Acetoclastic methanogenic bacteria (X4). The fifth step is where the Butyric acid (S6) is transformed into H2 (S3) and acetic acid (S5) by syntrophic bacteria B (X5). In the sixth step, both CO2 and H2 are used by the hydrogenotrophic-

methanogenic bacteria (X3) to generate methane (S7), and Transfer of CO2, H2 and methane between the gas and liquid phases of the bioreactor.

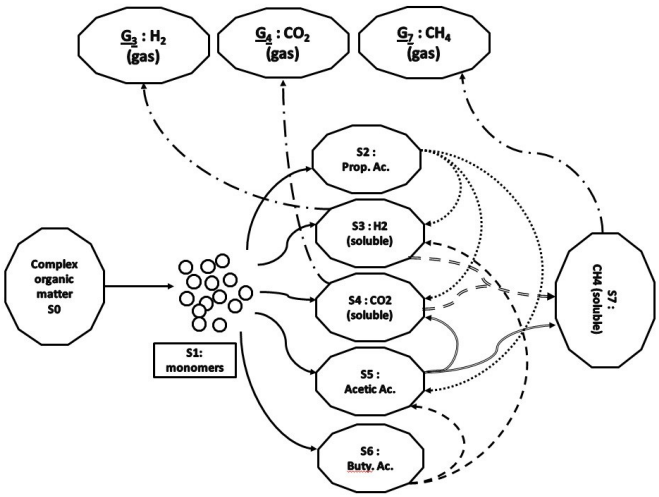


Fig 4. Diagram of main reactions and biochemical processes undertaken by the proposed model

IV. Model Parameter Identification Procedure

The model developed in the previous sections includes seventy-eight parameters that have to be identified from experimental data. The latter step is highly important to assure a wide range of validity of our model. For more practicability a structural identifiability problems, we have chosen an approach based on two points. We first decoupled the model parameters to be estimated into three groups: the kinetic parameters ($Kh, \mu_{MAX1}, K_{S_1}, \mu_{MAX2}, K_{S_2}, \mu_{MAX3}, K_{S_3}, \mu_{MAX4}, K_{S_4}, \mu_{MAX5}, K_{S_5}, K_{S_6}$), the transfer coefficient (Kla_3, Kla_4, Kla_7) and the yield coefficients. Our main motivation behind this decoupling lies in the high difficulty of kinetics modelling in general lead to a large uncertainty in bioprocess dynamical models. Secondly, we applied an optimisation algorithm to construct optimisation problems in MATLAB and solve them using the OPTimisation Interface (OPTI) toolbox[1].

V. Model Uncertainty and Sensitivity Analysis (UA)

Determining the strength of the relation between a given uncertain input and the output is the job of sensitivity analysis. For the latter reasons, we applied the US technique in this manuscript.

A. Uncertainty analysis (UA)

In this section, we explore the most popular sampling-based approaches used to perform UA, MC methods, and their most efficient implementation, namely the LHS technique.

[1] Copyright (C) 2011-2017, Jonathan Currie

1) Monte Carlo simulation (MC)

Random numbers are sampled from probability distributions of the model inputs. Each combination of model inputs is evaluated and the results can be used to both determine the uncertainty in model output and perform sensitivity analysis. A large number of samples are likely required to recreate input factor distributions through sampling. In the case where too little iteration is performed, not all values may be represented in the samples or values in the outer ranges may be under-sampled. To address the latter problem the Latin Hypercube Sampling (LHS) algorithm was specifically developed and it is by far the most popular sampling scheme for UA [14] .

2) Latin hypercube sampling (LHS)

The LHS technique allows an estimation of the average dynamic model outputs, with the benefit that it requires fewer samples than the classic random sampling to determine the same accuracy [15]. The random parameter distributions are divided into N (the sample size) equal probability intervals, which are then sampled. The sampling size N shall be at least $k+1$, where k is the number of parameters. The sampling is done by randomly selecting values from each probability density function (*pdf*). Each interval for each parameter is sampled exactly once (without replacement) so that the entire range for each parameter is explored. A matrix is generated (which we call the LHS matrix) that consists of N rows of the number of simulations (sample size) and of k columns corresponding to the number of varied parameters. N model solutions are then simulated, using each combination of parameter values (each row of the LHS matrix).

B. Sensitivity analysis (SA)

The application of sensitivity analysis will help determining the model critical parameters that control the model output. Global SA is an innovative approach for determining which reactions and processes contribute most to the behaviour of the overall system. Such techniques are usually processed using MC simulations and called sampling-based methods. Non-linearity between model inputs and outputs can be detected by using the scatter plot technique. The latter also helps the detection of non-monotonicities and correlations between them. For linear trends, the Pearson correlation coefficient (CC), partial correlation coefficients (PCCs), and standardised regression coefficients (SRC) work well in measuring the linearity between model inputs and outputs. For nonlinear but monotonic relationships between outputs and inputs, measures that work well are based on rank transforms such as Spearman rank correlation coefficient (RCC), partial rank correlation coefficient (PRCC), and standardised rank regression coefficients (SRRC). For nonlinear non-monotonic trends, methods based on the decomposition of model output variance are the best choice. The foremost method based on the variance decomposition is the Sobol method and its extended version using quasi-random numbers and an ad hoc design. We can also include the Fourier amplitude sensitivity test (FAST) and its extended version (eFAST). Both PRCC and eFAST measure two different model properties mainly the model monotonicity after the removal of the linear effects of all but one variable

and eFAST provides return measures of the variance fraction accounted for by individual variables and groups of variables. The Best-Case scenario is to use both of the two indexes to have a complete and informative US analysis.

VI. Discussion and conclusion

We applied the uncertainty and sensitivity (US) analyses to assess the adequacy of our dynamic AcoD complex model and establish which factors affect our model main outputs (G7(t): methane). A sampling-based method (partial rank correlation coefficient (PRCC)) and a variance-based method (extended Fourier amplitude sensitivity test—eFAST) were integrated to our US processing. The PRCC indexes provide answers to questions such as how the output is affected if we increase (or decrease) a specific parameter (linearly discounting the effects of the uncertainty over the rest of the parameters). Thus, the usage of PRCC reveals on what parameters to target if we want to achieve specific goals (e.g., control or regulation of the biogas production). For example, the most significant set of parameters can be used to determine how to efficiently reduce Volatile Fatty Acids (VFA) load or increase methane responses (by both timing and magnitude). Alternatively, eFAST as a variance-based method, in general, indicates which parameter uncertainty has the greatest impact on output variability.

Our results confirmed that our deterministic model simulation done after applying the parameter estimation algorithm was able to fit well with the observations (Fig.5).

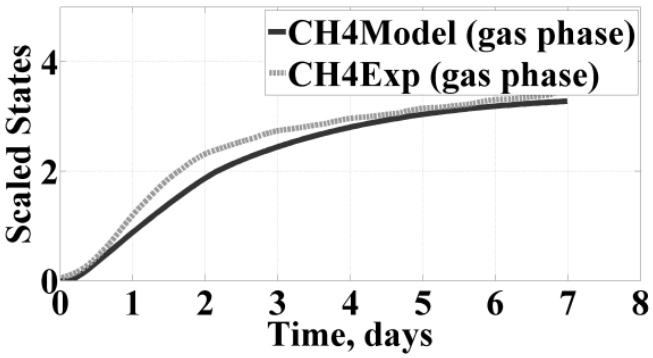
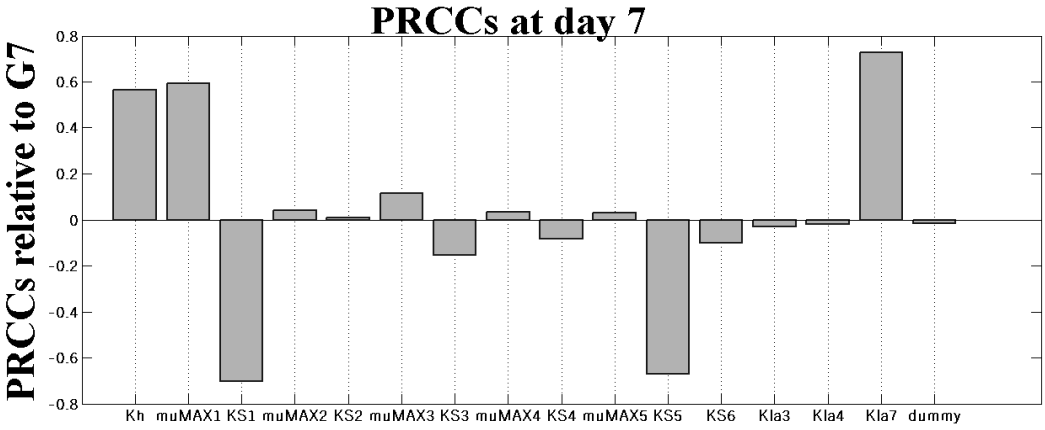


Fig 6. The Methane model estimation time series is validated against the observed data for mixed fruit, vegetables and cooked oil

Fig 5. PRCC performed on the AcoD complex model. The reference output is the methane gas production G7 at t = 7 days (end of experiment). PRCC results are calculated using a sample size N=1500.



After, the cumulative distribution functions (CDFs) of the parameter samples resulting from the LHS scheme were calculated. We set the sample size N to 1500. Each parameter is independently sampled from normal *pdfs* and the model was simulated for each parameter combination. The time point chosen to perform the SA is taken to be equal to the end of the duration of our experiments which is corresponding to the final methane production. Our UA demonstrated that PRCC provided nine significant ($p<0.01$) identified most influenced parameters on the model outputs (methane): Kh, μ_{MAX1}, K_{S_1}, μ_{MAX3} K_{S_3}, K_{S_4}, K_{S_5}, K_{S_6} and $Kla7$ (Fig. 6).

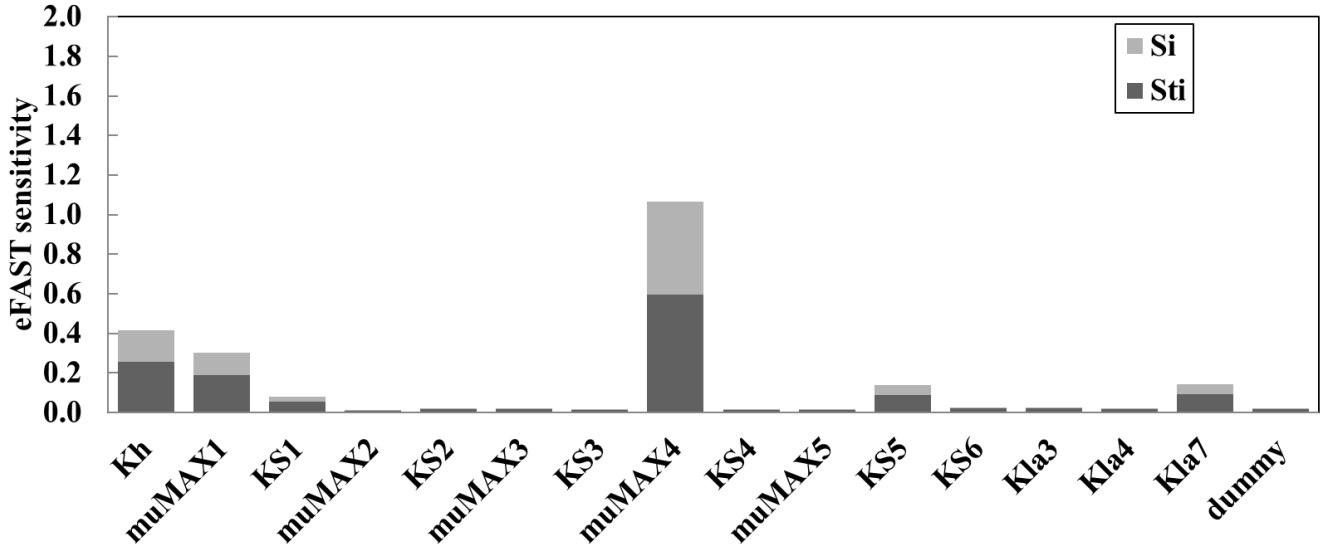


Fig 7. eFAST results with resampling and significance testing. Search curves were resampled five times (NR = 5), for a total of 1285 model evaluations.

First order *Si* and total order *Sti* are shown in Fig. 7 for each parameter, including a dummy parameter. Saltelli et al.[16], used dummy parameters as a screening method. The latter method inspired Marino et al. [17] in order to implement an equivalent technique in the context of eFAST with the purpose of testing the significance of first and total order sensitivity indexes.

Our analysis demonstrated that eFAST results demonstrated that the sets of significant parameters returned are mostly the same listed as significant by PRCC and are generally have smaller magnitude except the maximum acetoclastic methanogens (***muMAX4***). In fact, anaerobic methane production is very sensitive to acetate since they form about 70% of the methanogenic substrates in anaerobic digestors and is the only dicarbon substrate that methanogenic bacteria can degrade completely.

In conclusion, a complex AcoD model is proposed to simulate anaerobic co-digestion processes evolved in an anaerobic CSTR. A study of the aleatory uncertainty, a PRCC and eFAST analysis were done. The first step was to create parameter combinations using the LHS internal sampling algorithm. The second step was repeatedly executing the multiple complex AcoD model simulations for each parameter combination (replication step). Here it is worth noting that when using a pseudo-random number generator in the algorithm, we reinitialize the random seed for each model simulation. Finally, the sensitivity coefficients are calculated using the model outputs across replicates. The results obtained showed a good concordance between PRCC and eFAST. The most identified parameters that influence the most the methane production (model output G7) are $\boldsymbol{Kh}$, $\boldsymbol{\mu_{MAX1}}$, $\boldsymbol{K_{S_1}}$, $\boldsymbol{\mu_{MAX4}}$, $\boldsymbol{K_{S_5}}$ and $\boldsymbol{Kla_7}$. Thus, we were able to achieve our main goal which is having reliable information on which parameters to target if we want to control the model methane production. For example, the increasing of the polymer hydrolysis rate ($\boldsymbol{Kh}$), is shown to be effective on the increase of the final methane production. In addition, eFAST, indicate that ($\boldsymbol{Kh}$) uncertainty has the second greatest impact on the methane variability. We were able to simultaneously identify the parameters (i.e. biological mechanisms) that are driving or model outputs. These mechanisms can then be posed to the experimental community to execute more tests in a confidently. This close interaction between theorists and experimentalists provides the greatest opportunity for the use of mathematical models.

ACKNOWLEDGMENT

The laboratory experimental material used in this research work is a part of the funded research (Grant 63) from the " The Research Council of Oman" (TRC), The Sultanate of Oman.

REFERENCES

[1] H. Bouallagui, R. Ben Cheikh, L. Marouani, and M. Hamdi, "Bouallagui_2003_Mesophilic biogas production from fruit and vegetable waste in a tubular digester," vol. 86, pp. 85–89, 2003.

[2] H. Yazidi and J. V. Thanikal, "Nonlinear kinetic modelling of anaerobic biodegradation of fruit and vegetable waste together with cooked oil," Int. J. Adv. Res. Biol. Sci., vol. 3, no. 5, pp. 109–121, 2016.

[3] D. Dochain, Automatic control of bioprocess, vol. 53, no. 9. JohnWiley & Sons, Inc., ISTE Ltd., 2013.

[4] D. J. Hills and D. W. Roberts, "Anaerobic digestion of dairy manure and field crop residues," Agric. Wastes, vol. 3, no. 3, pp. 179–189, 1981.

[5] J. D. Bryers, "Structured modelling of the anaerobic digestion of biomass particles," Biotechnol. Bioeng., vol. 27, pp. 638–649, 1985.

[6] H. Siegrist, D. Renggli, and W. Gujer, "Mathematical modelling of anaerobic mesophilic sewage sludge treatment," in Water Science and Technology, 1993, vol. 27, no. 2, pp. 25–36.

[7] R. Moletta, D. Verrjer, and G. Albagnac, “Dynamic modelling of anaerobic digestion,” Water research, vol. 20, no. 4. pp. 427–434, 1986.
[8] V. A. Vavilin, S. V. Rytov, L. Y. Lokshina, J. A. Rintala, and G. Lyberatos, “Simplified hydrolysis models for the optimal design of two-stage anaerobic digestion,” Water Res., vol. 35, no. 17, pp. 4247–4251, 2001.
[9] D. J. Batstone et al., “The IWA Anaerobic Digestion Model No 1 (ADM1).,” Water Sci. Technol., vol. 45, no. 10, pp. 65–73, 2002.
[10] D. J. Batstone, J. Keller, and J. P. Steyer, “A review of ADM1 extensions, applications, and analysis: 2002-2005,” Water Science and Technology, vol. 54, no. 4. pp. 1–10, 2006.
[11] APHA, “Standard Methods for the Examination of Water and Wastewater,” Stand. Methods, p. 541, 2012.
[12] L. Appels et al., “Anaerobic digestion in global bio-energy production: Potential and research challenges,” Renew. Sustain. Energy Rev., vol. 15, no. 9, pp. 4295–4301, 2011.
[13] H. Yazidi and J. V. Thanikal, “Biodegradability and Bio Methane Potential of Vegetable, Fruit and Oil Fraction in Anaerobic Co,” Int. J. Curr. Res., vol. 7, no. 7, pp. 18379–18382, 2015.
[14] J. C. Helton, F. J. Davis, and J. D. Johnson, “A comparison of uncertainty and sensitivity analysis results obtained with random and Latin hypercube sampling,” Reliab. Eng. Syst. Saf., vol. 89, no. 3, pp. 305–330, 2005.
[15] M. D. Mckay, R. J. Beckman, and W. J. Conover, “A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code,” Technometrics, vol. 42, no. 1, p. 55, 2000.
[16] A. Saltelli, M. Ratto, S. Tarantola, and F. Campolongo, Sensitivity analysis practice: A guide to scientific models, vol. 91, no. 10–11. 2006.
[17] S. Marino, I. B. Hogue, C. J. Ray, and D. E. Kirschner, “A methodology for performing global uncertainty and sensitivity analysis in systems biology,” J. Theor. Biol., vol. 254, no. 1, pp. 178–196, 2008.